



33  
- 34 -

# SEQUENCE LISTING

## GENERAL INFORMATION:

- T340X
- AM
- (i) APPLICANT: Amara, Susan G  
Arriza, Jeffrey L
  - (ii) TITLE OF INVENTION: Amino Acid Transporters and Uses
  - (iii) NUMBER OF SEQUENCES: 17
  - (iv) CORRESPONDENCE ADDRESS:
    - (A) ADDRESSEE: Allegretti & Witcoff, Ltd.
    - (B) STREET: 10 South Wacker Drive, Suite 3000
    - (C) CITY: Chicago
    - (D) STATE: IL
    - (E) COUNTRY: USA
    - (F) ZIP: 60606
  - (v) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk
    - (B) COMPUTER: IBM PC compatible
    - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
  - (vi) CURRENT APPLICATION DATA:
    - (A) APPLICATION NUMBER: US 08/140,729
    - (B) FILING DATE: 20 OCT 1993
    - (C) CLASSIFICATION:
  - (viii) ATTORNEY/AGENT INFORMATION:
    - (A) NAME: Noonan, Kevin E
    - (B) REGISTRATION NUMBER: 35,303
    - (C) REFERENCE/DOCKET NUMBER: 93,509
  - (ix) TELECOMMUNICATION INFORMATION:
    - (A) TELEPHONE: 312-715-1000
    - (B) TELEFAX: 312-715-1234
    - (C) TELEX: 910-221-5317

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 63 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTGRGCRATG AARATGGCAG CCAGGGCYTC ATACAGGGCT GTGCCRTCCA TGTRATGGT 60  
RGC 63

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1680 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: 5'UTR  
(B) LOCATION: 1..30

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 31..1626

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR  
(B) LOCATION: 1626..1680

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CACCTCTAGC TCGGAGCGGC GTGTAGCGCC ATG GAG AAG AGC AAC GAG ACC AAC 54  
Met Glu Lys Ser Asn Glu Thr Asn  
1 5

GGC TAC CTT GAC AGC GCT CAG GCG GGG CCT GCG GCC GGG CCC GGA GCT 102  
Gly Tyr Leu Asp Ser Ala Gln Ala Gly Pro Ala Ala Gly Pro Gly Ala  
10 15 20

CCG GGG ACC GCG GCG GGA CGC GCA CGG CGT TGC GCG CGC TTC CTG CGG 150  
Pro Gly Thr Ala Ala Gly Arg Ala Arg Arg Cys Ala Arg Phe Leu Arg  
25 30 35 40

CGC CAA GCG CTG GTG CTG CTC ACC GTG TCC GGG GTG CTG GCG GGC GCG 198  
Arg Gln Ala Leu Val Leu Leu Thr Val Ser Gly Val Leu Ala Gly Ala  
45 50 55

GGC CTG GGC GCG GCG TTG CGC GGG CTC AGC CTG AGC CGC ACG CAG GTC 246  
Gly Leu Gly Ala Ala Leu Arg Gly Leu Ser Leu Ser Arg Thr Gln Val  
60 65 70

35  
-36-

ACC TAC CTG GCC TTC CCC GGC GAG ATG CTG CTC CGC ATG CTG CGC ATG	294
Thr Tyr Leu Ala Phe Pro Gly Glu Met Leu Leu Arg Met Leu Arg Met	
75 80 85	
ATC ATC CTG CCG CTG GTG GTC TGC AGC CTG GTG TCG GGC GCC GCC TCG	342
Ile Ile Leu Pro Leu Val Val Cys Ser Leu Val Ser Gly Ala Ala Ser	
90 95 100	
CTC GAT GCC AGC TGC CTC GGG CGT CTG GGC GGC ATC CGT GTC GCC TAC	390
Leu Asp Ala Ser Cys Leu Gly Arg Leu Gly Gly Ile Arg Val Ala Tyr	
105 110 115 120	
TTT GGC CTC ACC ACA CTG AGT GCC TCG GCG CTC GCC GTG GCC TTG GCG	438
Phe Gly Leu Thr Thr Leu Ser Ala Ser Ala Leu Ala Val Ala Leu Ala	
125 130 135	
TTC ATC ATC AAG CCA GGA TCC GGT GCG CAG ACC CTT CAG TCC AGC GAC	486
Phe Ile Ile Lys Pro Gly Ser Gly Ala Gln Thr Leu Gln Ser Ser Asp	
140 145 150	
CTG GGG CTG GAG GAC TCG GGG CCT CCT CCT GTC CCC AAA GAG ACG GTG	534
Leu Gly Leu Glu Asp Ser Gly Pro Pro Pro Val Pro Lys Glu Thr Val	
155 160 165	
GAC TCT TTC CTC GAC CTG GCC AGA AAC CTG TTT CCC TCC AAT CTT GTG	582
Asp Ser Phe Leu Asp Leu Ala Arg Asn Leu Phe Pro Ser Asn Leu Val	
170 175 180	
GTT GCA GCT TTC CGT ACG TAT GCA ACC GAT TAT AAA GTC GTG ACC CAG	630
Val Ala Ala Phe Arg Thr Tyr Ala Thr Asp Tyr Lys Val Val Thr Gln	
185 190 195 200	
AAC AGC AGC TCT GGA AAT GTA ACC CAT GAA AAG ATC CCC ATA GGC ACT	678
Asn Ser Ser Ser Gly Asn Val Thr His Glu Lys Ile Pro Ile Gly Thr	
205 210 215	
GAG ATA GAA GGG ATG AAC ATT TTA GGA TTG GTC CTG TTT GCT CTG GTG	726
Glu Ile Glu Gly Met Asn Ile Leu Gly Leu Val Leu Phe Ala Leu Val	
220 225 230	
TTA GGA GTG GCC TTA AAG AAA CTA GGC TCC GAA GGA GAA GAC CTC ATC	774
Leu Gly Val Ala Leu Lys Lys Leu Gly Ser Glu Gly Glu Asp Leu Ile	
235 240 245	
CGT TTC TTC AAT TCC CTC AAC GAG GCG ACG ATG GTG CTG GTG TCC TGG	822
Arg Phe Phe Asn Ser Leu Asn Glu Ala Thr Met Val Leu Val Ser Trp	
250 255 260	
ATT ATG TGG TAC GTA CCT GTG GGC ATC ATG TTC CTT GTT GGA AGC AAG	870
Ile Met Trp Tyr Val Pro Val Gly Ile Met Phe Leu Val Gly Ser Lys	
265 270 275 280	

36

ATC	GTG	GAA	ATG	AAA	GAC	ATC	ATC	GTG	CTG	GTG	ACC	AGC	CTG	GGG	AAA	918
Ile	Val	Glu	Met	Lys	Asp	Ile	Ile	Val	Leu	Val	Thr	Ser	Leu	Gly	Lys	
				285					290					295		
TAC	ATC	TTC	GCA	TCT	ATA	TTG	GGC	CAT	GTT	ATT	CAT	GGA	GGA	ATT	GTT	966
Tyr	Ile	Phe	Ala	Ser	Ile	Leu	Gly	His	Val	Ile	His	Gly	Gly	Ile	Val	
			300					305						310		
CTG	CCA	CTT	ATT	TAT	TTT	GTT	TTC	ACA	CGA	AAA	AAC	CCA	TTC	AGA	TTC	1014
Leu	Pro	Leu	Ile	Tyr	Phe	Val	Phe	Thr	Arg	Lys	Asn	Pro	Phe	Arg	Phe	
			315					320						325		
CTC	CTG	GGC	CTC	CTC	GCC	CCA	TTT	GCG	ACA	GCA	TTT	GCT	ACC	TGC	TCC	1062
Leu	Leu	Gly	Leu	Leu	Ala	Pro	Phe	Ala	Thr	Ala	Phe	Ala	Thr	Cys	Ser	
			330					335						340		
AGC	TCA	GCG	ACC	CTT	CCC	TCT	ATG	ATG	AAG	TGC	ATT	GAA	GAG	AAC	AAT	1110
Ser	Ser	Ala	Thr	Leu	Pro	Ser	Met	Met	Lys	Cys	Ile	Glu	Glu	Asn	Asn	
					345				350					355		360
GGT	GTG	GAC	AAG	AGG	ATC	AGC	AGG	TTT	ATT	CTC	CCC	ATC	GGG	GCC	ACC	1158
Gly	Val	Asp	Lys	Arg	Ile	Ser	Arg	Phe	Ile	Leu	Pro	Ile	Gly	Ala	Thr	
				365					370					375		
GTG	AAC	ATG	GAC	GGA	GCA	GCC	ATC	TTC	CAG	TGT	GTG	GCC	GCG	GTG	TTC	1206
Val	Asn	Met	Asp	Gly	Ala	Ala	Ile	Phe	Gln	Cys	Val	Ala	Ala	Val	Phe	
				380					385					390		
ATT	GCG	CAA	CTC	AAC	AAC	ATA	GAG	CTC	AAC	GCA	GGA	CAG	ATT	TTC	ACC	1254
Ile	Ala	Gln	Leu	Asn	Asn	Ile	Glu	Leu	Asn	Ala	Gly	Gln	Ile	Phe	Thr	
				395					400					405		
ATT	CTA	GTG	ACT	GCC	ACA	GCG	TCC	AGT	GTT	GGA	GCA	GCA	GGC	GTG	CCA	1302
Ile	Leu	Val	Thr	Ala	Thr	Ala	Ser	Ser	Val	Gly	Ala	Ala	Gly	Val	Pro	
				410					415					420		
GCT	GGA	GGG	GTC	CTC	ACC	ATT	GCC	ATT	ATC	CTG	GAG	GCC	ATT	GGG	CTG	1350
Ala	Gly	Gly	Val	Leu	Thr	Ile	Ala	Ile	Ile	Leu	Glu	Ala	Ile	Gly	Leu	
				425					430					435		440
CCT	ACT	CAT	GAC	CTG	CCT	CTG	ATC	CTG	GCT	GTG	GAC	TGG	ATT	GTG	GAC	1398
Pro	Thr	His	Asp	Leu	Pro	Leu	Ile	Leu	Ala	Val	Asp	Trp	Ile	Val	Asp	
				445					450					455		
CGG	ACC	ACC	ACG	GTG	GTG	AAT	GTG	GAG	GGG	GAT	GCC	CTG	GGT	GCA	GGC	1446
Arg	Thr	Thr	Thr	Val	Val	Asn	Val	Glu	Gly	Asp	Ala	Leu	Gly	Ala	Gly	
				460					465					470		
ATT	CTC	CAC	CAC	CTG	AAT	CAG	AAG	GCA	ACA	AAG	AAA	GGC	GAG	CAG	GAA	1494
Ile	Leu	His	His	Leu	Asn	Gln	Lys	Ala	Thr	Lys	Lys	Gly	Glu	Gln	Glu	
				475					480					485		

CTT GCT GAG GTG AAA GTG GAA GCC ATC CCC AAC TGC AAG TCT GAG GAG Leu Ala Glu Val Lys Val Glu Ala Ile Pro Asn Cys Lys Ser Glu Glu 490 495 500	1542
GAG ACA TCG CCC CTG GTG ACA CAC CAG AAC CCC GCT GGC CCC GTG GCC Glu Thr Ser Pro Leu Val Thr His Gln Asn Pro Ala Gly Pro Val Ala 505 510 515 520	1590
AGT GCC CCA GAA CTG GAA TCC AAG GAG TCG GTT CTG TGATGGGGCT Ser Ala Pro Glu Leu Glu Ser Lys Glu Ser Val Leu 525 530	1636
GGGCTTTGGG CTTGCCTGCC AGCAGTGATG TCCCACCCTG TTCA	1680

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 532 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Glu Lys Ser Asn Glu Thr Asn Gly Tyr Leu Asp Ser Ala Gln Ala 1 5 10 15
Gly Pro Ala Ala Gly Pro Gly Ala Pro Gly Thr Ala Ala Gly Arg Ala 20 25 30
Arg Arg Cys Ala Arg Phe Leu Arg Arg Gln Ala Leu Val Leu Leu Thr 35 40 45
Val Ser Gly Val Leu Ala Gly Ala Gly Leu Gly Ala Ala Leu Arg Gly 50 55 60
Leu Ser Leu Ser Arg Thr Gln Val Thr Tyr Leu Ala Phe Pro Gly Glu 65 70 75 80
Met Leu Leu Arg Met Leu Arg Met Ile Ile Leu Pro Leu Val Val Cys 85 90 95
Ser Leu Val Ser Gly Ala Ala Ser Leu Asp Ala Ser Cys Leu Gly Arg 100 105 110
Leu Gly Gly Ile Arg Val Ala Tyr Phe Gly Leu Thr Thr Leu Ser Ala 115 120 125
Ser Ala Leu Ala Val Ala Leu Ala Phe Ile Ile Lys Pro Gly Ser Gly 130 135 140

Ala Gln Thr Leu Gln Ser Ser Asp Leu Gly Leu Glu Asp Ser Gly Pro  
 145 150 155 160

Pro Pro Val Pro Lys Glu Thr Val Asp Ser Phe Leu Asp Leu Ala Arg  
 165 170 175

Asn Leu Phe Pro Ser Asn Leu Val Val Ala Ala Phe Arg Thr Tyr Ala  
 180 185 190

Thr Asp Tyr Lys Val Val Thr Gln Asn Ser Ser Ser Gly Asn Val Thr  
 195 200 205

His Glu Lys Ile Pro Ile Gly Thr Glu Ile Glu Gly Met Asn Ile Leu  
 210 215 220

Gly Leu Val Leu Phe Ala Leu Val Leu Gly Val Ala Leu Lys Lys Leu  
 225 230 235 240

Gly Ser Glu Gly Glu Asp Leu Ile Arg Phe Phe Asn Ser Leu Asn Glu  
 245 250 255

Ala Thr Met Val Leu Val Ser Trp Ile Met Trp Tyr Val Pro Val Gly  
 260 265 270

Ile Met Phe Leu Val Gly Ser Lys Ile Val Glu Met Lys Asp Ile Ile  
 275 280 285

Val Leu Val Thr Ser Leu Gly Lys Tyr Ile Phe Ala Ser Ile Leu Gly  
 290 295 300

His Val Ile His Gly Gly Ile Val Leu Pro Leu Ile Tyr Phe Val Phe  
 305 310 315 320

Thr Arg Lys Asn Pro Phe Arg Phe Leu Leu Gly Leu Leu Ala Pro Phe  
 325 330 335

Ala Thr Ala Phe Ala Thr Cys Ser Ser Ser Ala Thr Leu Pro Ser Met  
 340 345 350

Met Lys Cys Ile Glu Glu Asn Asn Gly Val Asp Lys Arg Ile Ser Arg  
 355 360 365

Phe Ile Leu Pro Ile Gly Ala Thr Val Asn Met Asp Gly Ala Ala Ile  
 370 375 380

Phe Gln Cys Val Ala Ala Val Phe Ile Ala Gln Leu Asn Asn Ile Glu  
 385 390 395 400

Leu Asn Ala Gly Gln Ile Phe Thr Ile Leu Val Thr Ala Thr Ala Ser  
 405 410 415

39  
-40-

Ser Val Gly Ala Ala Gly Val Pro Ala Gly Gly Val Leu Thr Ile Ala  
420 425 430

Ile Ile Leu Glu Ala Ile Gly Leu Pro Thr His Asp Leu Pro Leu Ile  
435 440 445

Leu Ala Val Asp Trp Ile Val Asp Arg Thr Thr Thr Val Val Asn Val  
450 455 460

Glu Gly Asp Ala Leu Gly Ala Gly Ile Leu His His Leu Asn Gln Lys  
465 470 475 480

Ala Thr Lys Lys Gly Glu Gln Glu Leu Ala Glu Val Lys Val Glu Ala  
485 490 495

Ile Pro Asn Cys Lys Ser Glu Glu Glu Thr Ser Pro Leu Val Thr His  
500 505 510

Gln Asn Pro Ala Gly Pro Val Ala Ser Ala Pro Glu Leu Glu Ser Lys  
515 520 525

Glu Ser Val Leu  
530

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1680 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: 5'UTR
- (B) LOCATION: 1..30

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 31..1656

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 1657..1680

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AAAGAAGAGA CCCTCCTAGA AAAGTAAAT ATG ACT AAA AGC AAT GGA GAA GAG  
Met Thr Lys Ser Asn Gly Glu Glu  
1 5

54

40

CCC AAG ATG GGG GGC AGG ATG GAG AGA TTC CAG CAG GGA GTC CGT AAA Pro Lys Met Gly Gly Arg Met Glu Arg Phe Gln Gln Gly Val Arg Lys 10 15 20	102
CGC ACA CTT TTG GCC AAG AAG AAA GTG CAG AAC ATT ACA AAG GAG GTT Arg Thr Leu Leu Ala Lys Lys Lys Val Gln Asn Ile Thr Lys Glu Val 25 30 35 40	150
GTT AAA AGT TAC CTG TTT CGG AAT GCT TTT GTG CTG CTC ACA GTC ACC Val Lys Ser Tyr Leu Phe Arg Asn Ala Phe Val Leu Leu Thr Val Thr 45 50 55	198
GCT GTC ATT GTG GGT ACA ATC CTT GGA TTT ACC CTC CGA CCA TAC AGA Ala Val Ile Val Gly Thr Ile Leu Gly Phe Thr Leu Arg Pro Tyr Arg 60 65 70	246
ATG AGC TAC CGG GAA GTC AAG TAC TTC TCC TTT CCT GGG GAA CTT CTG Met Ser Tyr Arg Glu Val Lys Tyr Phe Ser Phe Pro Gly Glu Leu Leu 75 80 85	294
ATG AGG ATG TTA CAG ATG CTG GTC TTA CCA CTT ATC ATC TCC AGT CTT Met Arg Met Leu Gln Met Leu Val Leu Pro Leu Ile Ile Ser Ser Leu 90 95 100	342
GTC ACA GGA ATG GCG GCG CTA GAT AGT AAG GCA TCA GGG AAG TGG GAA Val Thr Gly Met Ala Ala Leu Asp Ser Lys Ala Ser Gly Lys Trp Glu 105 110 115 120	390
TGC GGA GCT GTA GTC TAT TAT ATG ACT ACC ACC ATC ATT GCT GTG GTG Cys Gly Ala Val Val Tyr Tyr Met Thr Thr Thr Ile Ile Ala Val Val 125 130 135	438
ATT GGC ATA ATC ATT GTC ATC ATC ATC CAT CCT GGG AAG GGC ACA AAG Ile Gly Ile Ile Ile Val Ile Ile Ile His Pro Gly Lys Gly Thr Lys 140 145 150	486
GAA AAC ATG CAC AGA GAA GGC AAA ATT GTA CGA GTG ACA GCT GCA GAT Glu Asn Met His Arg Glu Gly Lys Ile Val Arg Val Thr Ala Ala Asp 155 160 165	534
GCC TTC CTG GAC TTG ATC AGG AAC ATG TTA AAT CCA AAT CTG GTA GAA Ala Phe Leu Asp Leu Ile Arg Asn Met Leu Asn Pro Asn Leu Val Glu 170 175 180	582
GCC TGC TTT AAA CAG TTT AAA ACC AAC TAT GAG AAG AGA AGC TTT AAA Ala Cys Phe Lys Gln Phe Lys Thr Asn Tyr Glu Lys Arg Ser Phe Lys 185 190 195 200	630
GTG CCC ATC CAG GCC AAC GAA ACG CTT GTG GGT GCT GTG ATA AAC AAT Val Pro Ile Gln Ala Asn Glu Thr Leu Val Gly Ala Val Ile Asn Asn 205 210 215	678



GTG TCT GAG GCC ATG GAG ACT CTT ACC CGA ATC ACA GAG GAG CTG GTC Val Ser Glu Ala Met Glu Thr Leu Thr Arg Ile Thr Glu Glu Leu Val 220 225 230	726
CCA GTT CCA GGA TCT GTG AAT GGA GTC AAT GCC CTG GGT CTA GTT GTC Pro Val Pro Gly Ser Val Asn Gly Val Asn Ala Leu Gly Leu Val Val 235 240 245	774
TTC TCC ATG TGC TTC GGT TTT GTG ATT GGA AAC ATG AAG GAA CAG GGG Phe Ser Met Cys Phe Gly Phe Val Ile Gly Asn Met Lys Glu Gln Gly 250 255 260	822
CAG GCC CTG AGA GAG TTC TTT GAT TCT CTT AAC GAA GCC ATC ATG AGA Gln Ala Leu Arg Glu Phe Phe Asp Ser Leu Asn Glu Ala Ile Met Arg 265 270 275 280	870
CTG GTA GCA GTA ATA ATG TGG TAT GCC CCC GTG GGT ATT CTC TTC CTG Leu Val Ala Val Ile Met Trp Tyr Ala Pro Val Gly Ile Leu Phe Leu 285 290 295	918
ATT GCT GGG AAG ATT GTG GAG ATG GAA GAC ATG GGT GTG ATT GGG GGG Ile Ala Gly Lys Ile Val Glu Met Glu Asp Met Gly Val Ile Gly Gly 300 305 310	966
CAG CTT GCC ATG TAC ACC GTG ACT GTC ATT GTT GGC TTA CTC ATT CAC Gln Leu Ala Met Tyr Thr Val Thr Val Ile Val Gly Leu Leu Ile His 315 320 325	1014
GCA GTC ATC GTC TTG CCA CTC CTC TAC TTC TTG GTA ACA CGG AAA AAC Ala Val Ile Val Leu Pro Leu Leu Tyr Phe Leu Val Thr Arg Lys Asn 330 335 340	1062
CCT TGG GTT TTT ATT GGA GGG TTG CTG CAA GCA CTC ATC ACC GCT CTG Pro Trp Val Phe Ile Gly Gly Leu Leu Gln Ala Leu Ile Thr Ala Leu 345 350 355 360	1110
GGG ACC TCT TCA AGT TCT GCC ACC CTA CCC ATC ACC TTC AAG TGC CTG Gly Thr Ser Ser Ser Ser Ala Thr Leu Pro Ile Thr Phe Lys Cys Leu 365 370 375	1158
GAA GAG AAC AAT GGC GTG GAC AAG CGC GTC ACC AGA TTC GTG CTC CCC Glu Glu Asn Asn Gly Val Asp Lys Arg Val Thr Arg Phe Val Leu Pro 380 385 390	1206
GTA GGA GCC ACC ATT AAC ATG GAT GGG ACT GCC CTC TAT GAG GCT TTG Val Gly Ala Thr Ile Asn Met Asp Gly Thr Ala Leu Tyr Glu Ala Leu 395 400 405	1254
GCT GCC ATT TTC ATT GCT CAA GTT AAC AAC TTT GAA CTG AAC TTC GGA Ala Ala Ile Phe Ile Ala Gln Val Asn Asn Phe Glu Leu Asn Phe Gly 410 415 420	1302

[illegible]

(2) INFORMATION FOR SEQ ID NO:5:

(i) **SEQUENCE CHARACTERISTICS:**

- (A) LENGTH: 542 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Thr	Lys	Ser	Asn	Gly	Glu	Glu	Pro	Lys	Met	Gly	Gly	Arg	Met	Glu
1				5					10					15	
Arg	Phe	Gln	Gln	Gly	Val	Arg	Lys	Arg	Thr	Leu	Leu	Ala	Lys	Lys	Lys
			20					25					30		
Val	Gln	Asn	Ile	Thr	Lys	Glu	Val	Val	Lys	Ser	Tyr	Leu	Phe	Arg	Asn
		35					40					45			

Ala	Phe	Val	Leu	Leu	Thr	Val	Thr	Ala	Val	Ile	Val	Gly	Thr	Ile	Leu	50	55	60	
Gly	Phe	Thr	Leu	Arg	Pro	Tyr	Arg	Met	Ser	Tyr	Arg	Glu	Val	Lys	Tyr	65	70	75	80
Phe	Ser	Phe	Pro	Gly	Glu	Leu	Leu	Met	Arg	Met	Leu	Gln	Met	Leu	Val	85	90	95	
Leu	Pro	Leu	Ile	Ile	Ser	Ser	Leu	Val	Thr	Gly	Met	Ala	Ala	Leu	Asp	100	105	110	
Ser	Lys	Ala	Ser	Gly	Lys	Trp	Glu	Cys	Gly	Ala	Val	Val	Tyr	Tyr	Met	115	120	125	
Thr	Thr	Thr	Ile	Ile	Ala	Val	Val	Ile	Gly	Ile	Ile	Ile	Val	Ile	Ile	130	135	140	
Ile	His	Pro	Gly	Lys	Gly	Thr	Lys	Glu	Asn	Met	His	Arg	Glu	Gly	Lys	145	150	155	160
Ile	Val	Arg	Val	Thr	Ala	Ala	Asp	Ala	Phe	Leu	Asp	Leu	Ile	Arg	Asn	165	170	175	
Met	Leu	Asn	Pro	Asn	Leu	Val	Glu	Ala	Cys	Phe	Lys	Gln	Phe	Lys	Thr	180	185	190	
Asn	Tyr	Glu	Lys	Arg	Ser	Phe	Lys	Val	Pro	Ile	Gln	Ala	Asn	Glu	Thr	195	200	205	
Leu	Val	Gly	Ala	Val	Ile	Asn	Asn	Val	Ser	Glu	Ala	Met	Glu	Thr	Leu	210	215	220	
Thr	Arg	Ile	Thr	Glu	Glu	Leu	Val	Pro	Val	Pro	Gly	Ser	Val	Asn	Gly	225	230	235	240
Val	Asn	Ala	Leu	Gly	Leu	Val	Val	Phe	Ser	Met	Cys	Phe	Gly	Phe	Val	245	250	255	
Ile	Gly	Asn	Met	Lys	Glu	Gln	Gly	Gln	Ala	Leu	Arg	Glu	Phe	Phe	Asp	260	265	270	
Ser	Leu	Asn	Glu	Ala	Ile	Met	Arg	Leu	Val	Ala	Val	Ile	Met	Trp	Tyr	275	280	285	
Ala	Pro	Val	Gly	Ile	Leu	Phe	Leu	Ile	Ala	Gly	Lys	Ile	Val	Glu	Met	290	295	300	
Glu	Asp	Met	Gly	Val	Ile	Gly	Gly	Gln	Leu	Ala	Met	Tyr	Thr	Val	Thr	305	310	315	320

Val Ile Val Gly Leu Leu Ile His Ala Val Ile Val Leu Pro Leu Leu  
325 330 335

Tyr Phe Leu Val Thr Arg Lys Asn Pro Trp Val Phe Ile Gly Gly Leu  
340 345 350

Leu Gln Ala Leu Ile Thr Ala Leu Gly Thr Ser Ser Ser Ser Ala Thr  
355 360 365

Leu Pro Ile Thr Phe Lys Cys Leu Glu Glu Asn Asn Gly Val Asp Lys  
370 375 380

Arg Val Thr Arg Phe Val Leu Pro Val Gly Ala Thr Ile Asn Met Asp  
385 390 395 400

Gly Thr Ala Leu Tyr Glu Ala Leu Ala Ala Ile Phe Ile Ala Gln Val  
405 410 415

Asn Asn Phe Glu Leu Asn Phe Gly Gln Ile Ile Thr Ile Ser Ile Thr  
420 425 430

Ala Thr Ala Ala Ser Ile Gly Ala Ala Gly Ile Pro Gln Ala Gly Leu  
435 440 445

Val Thr Met Val Ile Val Leu Thr Ser Val Gly Leu Pro Thr Asp Asp  
450 455 460

Ile Thr Leu Ile Ile Ala Val Asp Trp Phe Leu Asp Arg Leu Arg Thr  
465 470 475 480

Thr Thr Asn Val Leu Gly Asp Ser Leu Gly Ala Gly Ile Val Glu His  
485 490 495

Leu Ser Arg His Glu Leu Lys Asn Arg Asp Val Glu Met Gly Asn Ser  
500 505 510

Val Ile Glu Glu Asn Glu Met Lys Lys Pro Tyr Gln Leu Ile Ala Gln  
515 520 525

Asp Asn Glu Thr Glu Lys Pro Ile Asp Ser Glu Thr Lys Met  
530 535 540

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1800 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: 5'UTR

(B) LOCATION: 1..33

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 34..1755

(ix) FEATURE:

(A) NAME/KEY: 3'UTR

(B) LOCATION: 1756..1800

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GATAGTGCTG AAGAGGAGGG GCGTTCCTCAG ACC ATG GCA TCT ACG GAA GGT GCC	54
Met Ala Ser Thr Glu Gly Ala	
1 5	
AAC AAT ATG CCC AAG CAG GTG GAA GTG CGA ATG CCA GAC AGT CAT CTT	102
Asn Asn Met Pro Lys Gln Val Glu Val Arg Met Pro Asp Ser His Leu	
10 15 20	
GGC TCA GAG GAA CCC AAG CAC CGG CAC CTG GGC CTG CGC CTG TGT GAC	150
Gly Ser Glu Glu Pro Lys His Arg His Leu Gly Leu Arg Leu Cys Asp	
25 30 35	
AAG CTG GGG AAG AAT CTG CTG CTC ACC CTG ACG GTG TTT GGT GTC ATC	198
Lys Leu Gly Lys Asn Leu Leu Leu Thr Leu Thr Val Phe Gly Val Ile	
40 45 50 55	
CTG GGA GCA GTG TGT GGA GGG CTT CTT CGC TTG GCA TCT CCC ATC CAC	246
Leu Gly Ala Val Cys Gly Gly Leu Leu Arg Leu Ala Ser Pro Ile His	
60 65 70	
CCT GAT GTG GTT ATG TTA ATA GCC TTC CCA GGG GAT ATA CTC ATG AGG	294
Pro Asp Val Val Met Leu Ile Ala Phe Pro Gly Asp Ile Leu Met Arg	
75 80 85	
ATG CTA AAA ATG CTC ATT CTG GGT CTA ATC ATC TCC AGC TTA ATC ACA	342
Met Leu Lys Met Leu Ile Leu Gly Leu Ile Ile Ser Ser Leu Ile Thr	
90 95 100	
GGG TTG TCA GGC CTG GAT GCT AAG GCT AGT GGC CGC TTG GGC ACG AGA	390
Gly Leu Ser Gly Leu Asp Ala Lys Ala Ser Gly Arg Leu Gly Thr Arg	
105 110 115	
GCC ATG GTG TAT TAC ATG TCC ACG ACC ATC ATT GCT GCA GTA CTG GGG	438
Ala Met Val Tyr Tyr Met Ser Thr Thr Ile Ile Ala Ala Val Leu Gly	
120 125 130 135	

GTC ATT CTG GTC TTG GCT ATC CAT CCA GGC AAT CCC AAG CTC AAG AAG	486
Val Ile Leu Val Leu Ala Ile His Pro Gly Asn Pro Lys Leu Lys Lys	
140 145 150	
CAG CTG GGG CCT GGG AAG AAG AAT GAT GAA GTG TCC AGC CTG GAT GCC	534
Gln Leu Gly Pro Gly Lys Lys Asn Asp Glu Val Ser Ser Leu Asp Ala	
155 160 165	
TTC CTG GAC CTT ATT CGA AAT CTC TTC CCT GAA AAC CTT GTC CAA GCC	582
Phe Leu Asp Leu Ile Arg Asn Leu Phe Pro Glu Asn Leu Val Gln Ala	
170 175 180	
TGC TTT CAA CAG ATT CAA ACA GTG ACG AAG AAA GTC CTG GTT GCA CCA	630
Cys Phe Gln Gln Ile Gln Thr Val Thr Lys Lys Val Leu Val Ala Pro	
185 190 195	
CCG CCA GAC GAG GAG GCC AAC GCA ACC AGC GCT GAA GTC TCT CTG TTG	678
Pro Pro Asp Glu Glu Ala Asn Ala Thr Ser Ala Glu Val Ser Leu Leu	
200 205 210 215	
AAC GAG ACT GTG ACT GAG GTG CCG GAG GAG ACT AAG ATG GTT ATC AAG	726
Asn Glu Thr Val Thr Glu Val Pro Glu Glu Thr Lys Met Val Ile Lys	
220 225 230	
AAG GGC CTG GAG TTC AAG GAT GGG ATG AAC GTC TTA GGT CTG ATA GGG	774
Lys Gly Leu Glu Phe Lys Asp Gly Met Asn Val Leu Gly Leu Ile Gly	
235 240 245	
TTT TTC ATT GCT TTT GGC ATC GCT ATG GGG AAG ATG GGA GAT CAG GCC	822
Phe Phe Ile Ala Phe Gly Ile Ala Met Gly Lys Met Gly Asp Gln Ala	
250 255 260	
AAG CTG ATG GTG GAT TTC TTC AAC ATT TTG AAT GAG ATT GTA ATG AAG	870
Lys Leu Met Val Asp Phe Phe Asn Ile Leu Asn Glu Ile Val Met Lys	
265 270 275	
TTA GTG ATC ATG ATC ATG TGG TAC TCT CCC CTG GGT ATC GCC TGC CTG	918
Leu Val Ile Met Ile Met Trp Tyr Ser Pro Leu Gly Ile Ala Cys Leu	
280 285 290 295	
ATC TGT GGA AAG ATC ATT GCA ATC AAG GAC TTA GAA GTG GTT GCT AGG	966
Ile Cys Gly Lys Ile Ile Ala Ile Lys Asp Leu Glu Val Val Ala Arg	
300 305 310	
CAA CTG GGG ATG TAC ATG GTA ACA GTG ATC ATA GGC CTC ATC ATC CAC	1014
Gln Leu Gly Met Tyr Met Val Thr Val Ile Ile Gly Leu Ile Ile His	
315 320 325	
GGG GGC ATC TTT CTC CCC TTG ATT TAC TTT GTA GTG ACC AGG AAA AAC	1062
Gly Gly Ile Phe Leu Pro Leu Ile Tyr Phe Val Val Thr Arg Lys Asn	
330 335 340	

CCC TTC TCC CTT TTT GCT GGC ATT TTC CAA GCT TGG ATC ACT GCC CTG Pro Phe Ser Leu Phe Ala Gly Ile Phe Gln Ala Trp Ile Thr Ala Leu 345 350 355	1110
GGC ACC GCT TCC AGT GCT GGA ACT TTG CCT GTC ACC TTT CGT TGC CTG Gly Thr Ala Ser Ser Ala Gly Thr Leu Pro Val Thr Phe Arg Cys Leu 360 365 370 375	1158
GAA GAA AAT CTG GGG ATT GAT AAG CGT GTG ACT AGA TTC GTC CTT CCT Glu Glu Asn Leu Gly Ile Asp Lys Arg Val Thr Arg Phe Val Leu Pro 380 385 390	1206
GTT GGA GCA ACC ATT AAC ATG GAT GGT ACA GCC CTT TAT GAA GCG GTG Val Gly Ala Thr Ile Asn Met Asp Gly Thr Ala Leu Tyr Glu Ala Val 395 400 405	1254
GCC GCC ATC TTT ATA GCC CAA ATG AAT GGT GTT GTC CTG GAT GGA GGA Ala Ala Ile Phe Ile Ala Gln Met Asn Gly Val Val Leu Asp Gly Gly 410 415 420	1302
CAG ATT GTG ACT GTA AGC CTC ACA GCC ACC CTG GCA AGC GTC GGC GCG Gln Ile Val Thr Val Ser Leu Thr Ala Thr Leu Ala Ser Val Gly Ala 425 430 435	1350
GCC AGT ATC CCC AGT GCC GGG CTG GTC ACC ATG CTC CTC ATT CTG ACA Ala Ser Ile Pro Ser Ala Gly Leu Val Thr Met Leu Leu Ile Leu Thr 440 445 450 455	1398
GCC GTG GGC CTG CCA ACA GAG GAC ATC AGC TTG CTG GTG GCT GTG GAC Ala Val Gly Leu Pro Thr Glu Asp Ile Ser Leu Leu Val Ala Val Asp 460 465 470	1446
TGG CTG CTG GAC AGG ATG AGA ACT TCA GTC AAT GTT GTG GGT GAC TCT Trp Leu Leu Asp Arg Met Arg Thr Ser Val Asn Val Val Gly Asp Ser 475 480 485	1494
TTT GGG GCT GGG ATA GTC TAT CAC CTC TCC AAG TCT GAG CTG GAT ACC Phe Gly Ala Gly Ile Val Tyr His Leu Ser Lys Ser Glu Leu Asp Thr 490 495 500	1542
ATT GAC TCC CAG CAT CGA GTG CAT GAA GAT ATT GAA ATG ACC AAG ACT Ile Asp Ser Gln His Arg Val His Glu Asp Ile Glu Met Thr Lys Thr 505 510 515	1590
CAA TCC ATT TAT GAT GAC ATG AAG AAC CAC AGG GAA AGC AAC TCT AAT Gln Ser Ile Tyr Asp Asp Met Lys Asn His Arg Glu Ser Asn Ser Asn 520 525 530 535	1638
CAA TGT GTC TAT GCT GCA CAC AAC TCT GTC ATA GTA GAT GAA TGC AAG Gln Cys Val Tyr Ala Ala His Asn Ser Val Ile Val Asp Glu Cys Lys 540 545 550	1686





Glu	Val	Ser	Ser	Leu	Asp	Ala	Phe	Leu	Asp	Leu	Ile	Arg	Asn	Leu	Phe	165	170	175
Pro	Glu	Asn	Leu	Val	Gln	Ala	Cys	Phe	Gln	Gln	Ile	Gln	Thr	Val	Thr	180	185	190
Lys	Lys	Val	Leu	Val	Ala	Pro	Pro	Pro	Asp	Glu	Glu	Ala	Asn	Ala	Thr	195	200	205
Ser	Ala	Glu	Val	Ser	Leu	Leu	Asn	Glu	Thr	Val	Thr	Glu	Val	Pro	Glu	210	215	220
Glu	Thr	Lys	Met	Val	Ile	Lys	Lys	Gly	Leu	Glu	Phe	Lys	Asp	Gly	Met	225	230	235
Asn	Val	Leu	Gly	Leu	Ile	Gly	Phe	Phe	Ile	Ala	Phe	Gly	Ile	Ala	Met	245	250	255
Gly	Lys	Met	Gly	Asp	Gln	Ala	Lys	Leu	Met	Val	Asp	Phe	Phe	Asn	Ile	260	265	270
Leu	Asn	Glu	Ile	Val	Met	Lys	Leu	Val	Ile	Met	Ile	Met	Trp	Tyr	Ser	275	280	285
Pro	Leu	Gly	Ile	Ala	Cys	Leu	Ile	Cys	Gly	Lys	Ile	Ile	Ala	Ile	Lys	290	295	300
Asp	Leu	Glu	Val	Val	Ala	Arg	Gln	Leu	Gly	Met	Tyr	Met	Val	Thr	Val	305	310	315
Ile	Ile	Gly	Leu	Ile	Ile	His	Gly	Gly	Ile	Phe	Leu	Pro	Leu	Ile	Tyr	325	330	335
Phe	Val	Val	Thr	Arg	Lys	Asn	Pro	Phe	Ser	Leu	Phe	Ala	Gly	Ile	Phe	340	345	350
Gln	Ala	Trp	Ile	Thr	Ala	Leu	Gly	Thr	Ala	Ser	Ser	Ala	Gly	Thr	Leu	355	360	365
Pro	Val	Thr	Phe	Arg	Cys	Leu	Glu	Glu	Asn	Leu	Gly	Ile	Asp	Lys	Arg	370	375	380
Val	Thr	Arg	Phe	Val	Leu	Pro	Val	Gly	Ala	Thr	Ile	Asn	Met	Asp	Gly	385	390	395
Thr	Ala	Leu	Tyr	Glu	Ala	Val	Ala	Ala	Ile	Phe	Ile	Ala	Gln	Met	Asn	405	410	415
Gly	Val	Val	Leu	Asp	Gly	Gly	Gln	Ile	Val	Thr	Val	Ser	Leu	Thr	Ala	420	425	430

-51-50-

Thr Leu Ala Ser Val Gly Ala Ala Ser Ile Pro Ser Ala Gly Leu Val  
435 440 445

Thr Met Leu Leu Ile Leu Thr Ala Val Gly Leu Pro Thr Glu Asp Ile  
450 455 460

Ser Leu Leu Val Ala Val Asp Trp Leu Leu Asp Arg Met Arg Thr Ser  
465 470 475 480

Val Asn Val Val Gly Asp Ser Phe Gly Ala Gly Ile Val Tyr His Leu  
485 490 495

Ser Lys Ser Glu Leu Asp Thr Ile Asp Ser Gln His Arg Val His Glu  
500 505 510

Asp Ile Glu Met Thr Lys Thr Gln Ser Ile Tyr Asp Asp Met Lys Asn  
515 520 525

His Arg Glu Ser Asn Ser Asn Gln Cys Val Tyr Ala Ala His Asn Ser  
530 535 540

Val Ile Val Asp Glu Cys Lys Val Thr Leu Ala Ala Asn Gly Lys Ser  
545 550 555 560

Ala Asp Cys Ser Val Glu Glu Glu Pro Trp Lys Arg Glu Lys  
565 570

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1674 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: 5'UTR
- (B) LOCATION: 1..15

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16..1590

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 1591..1674

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

57

57  
- 52 -

ATAGCGGCGA CAGCC ATG GGG AAA CCG GCG AGG AAA GGA TGC CCG AGT TGG	51
Met Gly Lys Pro Ala Arg Lys Gly Cys Pro Ser Trp	
1 5 10	
AAG CGC TTC CTG AAG AAT AAC TGG GTG TTG CTG TCC ACC GTG GCC GCG	99
Lys Arg Phe Leu Lys Asn Asn Trp Val Leu Leu Ser Thr Val Ala Ala	
15 20 25	
GTG GTG CTA GGC ATT ACC ACA GGA GTC TTG GTT CGA GAA CAC AGC AAC	147
Val Val Leu Gly Ile Thr Thr Gly Val Leu Val Arg Glu His Ser Asn	
30 35 40	
CTC TCA ACT CTA GAG AAA TTC TAC TTT GCT TTT CCT GGA GAA ATT CTA	195
Leu Ser Thr Leu Glu Lys Phe Tyr Phe Ala Phe Pro Gly Glu Ile Leu	
45 50 55 60	
ATG CGG ATG CTG AAA CTC ATC ATT TTG CCA TTA ATT ATA TCC AGC ATG	243
Met Arg Met Leu Lys Leu Ile Ile Leu Pro Leu Ile Ile Ser Ser Met	
65 70 75	
ATT ACA GGT GTT GCT GCA CTG GAT TCC AAC GTA TCC GGA AAA ATT GGT	291
Ile Thr Gly Val Ala Ala Leu Asp Ser Asn Val Ser Gly Lys Ile Gly	
80 85 90	
CTG CGC GCT GTC GTG TAT TAT TTC TGT ACC ACT CTC ATT GCT GTT ATT	339
Leu Arg Ala Val Val Tyr Tyr Phe Cys Thr Thr Leu Ile Ala Val Ile	
95 100 105	
CTA GGT ATT GTG CTG GTG GTG AGC ATC AAG CCT GGT GTC ACC CAG AAA	387
Leu Gly Ile Val Leu Val Val Ser Ile Lys Pro Gly Val Thr Gln Lys	
110 115 120	
GTG GGT GAA ATT GCG AGG ACA GGC AGC ACC CCT GAA GTC AGT ACG GTG	435
Val Gly Glu Ile Ala Arg Thr Gly Ser Thr Pro Glu Val Ser Thr Val	
125 130 135 140	
GAT GCC ATG TTA GAT CTC ATC AGG AAT ATG TTC CCT GAG AAT CTT GTC	483
Asp Ala Met Leu Asp Leu Ile Arg Asn Met Phe Pro Glu Asn Leu Val	
145 150 155	
CAG GCC TGT TTT CAG CAG TAC AAA ACT AAG CGT GAA GAA GTG AAG CCT	531
Gln Ala Cys Phe Gln Gln Tyr Lys Thr Lys Arg Glu Glu Val Lys Pro	
160 165 170	
CCC AGC GAT CCA GAG ATG AAC ATG ACA GAA GAG TCC TTC ACA GCT GTC	579
Pro Ser Asp Pro Glu Met Asn Met Thr Glu Glu Ser Phe Thr Ala Val	
175 180 185	
ATG ACA ACT GCA ATT TCC AAG AAC AAA ACA AAG GAA TAC AAA ATT GTT	627
Met Thr Thr Ala Ile Ser Lys Asn Lys Thr Lys Glu Tyr Lys Ile Val	
190 195 200	

52

- 52 -

GGC ATG TAT TCA GAT GGC ATA AAC GTC CTG GGC TTG ATT GTC TTT TGC Gly Met Tyr Ser Asp Gly Ile Asn Val Leu Gly Leu Ile Val Phe Cys 205 210 215 220	675
CTT GTC TTT GGA CTT GTC ATT GGA AAA ATG GGA GAA AAG GGA CAA ATT Leu Val Phe Gly Leu Val Ile Gly Lys Met Gly Glu Lys Gly Gln Ile 225 230 235	723
CTG GTG GAT TTC TTC AAT GCT TTG AGT GAT GCA ACC ATG AAA ATC GTT Leu Val Asp Phe Phe Asn Ala Leu Ser Asp Ala Thr Met Lys Ile Val 240 245 250	771
CAG ATC ATC ATG TGT TAT ATG CCA CTA GGT ATT TTG TTC CTG ATT GCT Gln Ile Ile Met Cys Tyr Met Pro Leu Gly Ile Leu Phe Leu Ile Ala 255 260 265	819
GGG AAG ATC ATA GAA GTT GAA GAC TGG GAA ATA TTC CGC AAG CTG GGC Gly Lys Ile Ile Glu Val Glu Asp Trp Glu Ile Phe Arg Lys Leu Gly 270 275 280	867
CTT TAC ATG GCC ACA GTC CTG ACT GGG CTT GCA ATC CAC TCC ATT GTA Leu Tyr Met Ala Thr Val Leu Thr Gly Leu Ala Ile His Ser Ile Val 285 290 295 300	915
ATT CTC CCG CTG ATA TAT TTC ATA GTC GTA CGA AAG AAC CCT TTC CGA Ile Leu Pro Leu Ile Tyr Phe Ile Val Val Arg Lys Asn Pro Phe Arg 305 310 315	963
TTT GCC ATG GGA ATG GCC CAG GCT CTC CTG ACA GCT CTC ATG ATC TCT Phe Ala Met Gly Met Ala Gln Ala Leu Leu Thr Ala Leu Met Ile Ser 320 325 330	1011
TCC AGT TCA GCA ACA CTG CCT GTC ACC TTC CGC TGT GCT GAA GAA AAT Ser Ser Ser Ala Thr Leu Pro Val Thr Phe Arg Cys Ala Glu Glu Asn 335 340 345	1059
AAC CAG GTG GAC AAG AGG ATC ACT CGA TTC GTG TTA CCC GTT GGT GCA Asn Gln Val Asp Lys Arg Ile Thr Arg Phe Val Leu Pro Val Gly Ala 350 355 360	1107
ACA ATC AAC ATG GAT GGG ACC GCG CTC TAT GAA GCA GTG GCA GCG GTG Thr Ile Asn Met Asp Gly Thr Ala Leu Tyr Glu Ala Val Ala Ala Val 365 370 375 380	1155
TTT ATT GCA CAG TTG AAT GAC CTG GAC TTG GGC ATT GGG CAG ATC ATC Phe Ile Ala Gln Leu Asn Asp Leu Asp Leu Gly Ile Gly Gln Ile Ile 385 390 395	1203
ACC ATC AGT ATC ACG GCC ACA TCT GCC AGC ATC GGA GCT GCT GGC GTG Thr Ile Ser Ile Thr Ala Thr Ser Ala Ser Ile Gly Ala Ala Gly Val 400 405 410	1251

53

-54- 53

CCC CAG GCT GGC CTG GTG ACC ATG GTG ATT GTG CTG AGT GCC GTG GGC Pro Gln Ala Gly Leu Val Thr Met Val Ile Val Leu Ser Ala Val Gly 415 420 425	1299
CTG CCC GCC GAG GAT GTC ACC CTG ATC ATT GCT GTC GAC TGG CTC CTG Leu Pro Ala Glu Asp Val Thr Leu Ile Ile Ala Val Asp Trp Leu Leu 430 435 440	1347
GAC CGG TTC AGG ACC ATG GTC AAC GTC CTT GGT GAT GCT TTT GGG ACG Asp Arg Phe Arg Thr Met Val Asn Val Leu Gly Asp Ala Phe Gly Thr 445 450 455 460	1395
GGC ATT GTG GAA AAG CTC TCC AAG AAG GAG CTG GAG CAG ATG GAT GTT Gly Ile Val Glu Lys Leu Ser Lys Lys Glu Leu Glu Gln Met Asp Val 465 470 475	1443
TCA TCT GAA GTC AAC ATT GTG AAT CCC TTT GCC TTG GAA TCC ACA ATC Ser Ser Glu Val Asn Ile Val Asn Pro Phe Ala Leu Glu Ser Thr Ile 480 485 490	1491
CTT GAC AAC GAA GAC TCA GAC ACC AAG AAG TCT TAT GTC AAT GGA GGC Leu Asp Asn Glu Asp Ser Asp Thr Lys Lys Ser Tyr Val Asn Gly Gly 495 500 505	1539
TTT GCA GTA GAC AAG TCT GAC ACC ATC TCA TTC ACC CAG ACC TCA CAG Phe Ala Val Asp Lys Ser Asp Thr Ile Ser Phe Thr Gln Thr Ser Gln 510 515 520	1587
TTC TAGGGCCCCCT GGCTGCAGAT GACTGGAAAC AAGGAAGGAC ATTTCGTGAG Phe 525	1640
AGTCATCTCA AACACGGCTT AAGGAAAAGA GAAA	1674

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 525 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Gly Lys Pro Ala Arg Lys Gly Cys Pro Ser Trp Lys Arg Phe Leu 1 5 10 15
Lys Asn Asn Trp Val Leu Leu Ser Thr Val Ala Ala Val Val Leu Gly 20 25 30

54

54  
-35-

Ile Thr Thr Gly Val Leu Val Arg Glu His Ser Asn Leu Ser Thr Leu  
35 40 45

Glu Lys Phe Tyr Phe Ala Phe Pro Gly Glu Ile Leu Met Arg Met Leu  
50 55 60

Lys Leu Ile Ile Leu Pro Leu Ile Ile Ser Ser Met Ile Thr Gly Val  
65 70 75 80

Ala Ala Leu Asp Ser Asn Val Ser Gly Lys Ile Gly Leu Arg Ala Val  
85 90 95

Val Tyr Tyr Phe Cys Thr Thr Leu Ile Ala Val Ile Leu Gly Ile Val  
100 105 110

Leu Val Val Ser Ile Lys Pro Gly Val Thr Gln Lys Val Gly Glu Ile  
115 120 125

Ala Arg Thr Gly Ser Thr Pro Glu Val Ser Thr Val Asp Ala Met Leu  
130 135 140

Asp Leu Ile Arg Asn Met Phe Pro Glu Asn Leu Val Gln Ala Cys Phe  
145 150 155 160

Gln Gln Tyr Lys Thr Lys Arg Glu Glu Val Lys Pro Pro Ser Asp Pro  
165 170 175

Glu Met Asn Met Thr Glu Glu Ser Phe Thr Ala Val Met Thr Thr Ala  
180 185 190

Ile Ser Lys Asn Lys Thr Lys Glu Tyr Lys Ile Val Gly Met Tyr Ser  
195 200 205

Asp Gly Ile Asn Val Leu Gly Leu Ile Val Phe Cys Leu Val Phe Gly  
210 215 220

Leu Val Ile Gly Lys Met Gly Glu Lys Gly Gln Ile Leu Val Asp Phe  
225 230 235 240

Phe Asn Ala Leu Ser Asp Ala Thr Met Lys Ile Val Gln Ile Ile Met  
245 250 255

Cys Tyr Met Pro Leu Gly Ile Leu Phe Leu Ile Ala Gly Lys Ile Ile  
260 265 270

Glu Val Glu Asp Trp Glu Ile Phe Arg Lys Leu Gly Leu Tyr Met Ala  
275 280 285

Thr Val Leu Thr Gly Leu Ala Ile His Ser Ile Val Ile Leu Pro Leu  
290 295 300

55

55  
-36-

Ile Tyr Phe Ile Val Val Arg Lys Asn Pro Phe Arg Phe Ala Met Gly  
305 310 315 320

Met Ala Gln Ala Leu Leu Thr Ala Leu Met Ile Ser Ser Ser Ser Ala  
325 330 335

Thr Leu Pro Val Thr Phe Arg Cys Ala Glu Glu Asn Asn Gln Val Asp  
340 345 350

Lys Arg Ile Thr Arg Phe Val Leu Pro Val Gly Ala Thr Ile Asn Met  
355 360 365

Asp Gly Thr Ala Leu Tyr Glu Ala Val Ala Ala Val Phe Ile Ala Gln  
370 375 380

Leu Asn Asp Leu Asp Leu Gly Ile Gly Gln Ile Ile Thr Ile Ser Ile  
385 390 395 400

Thr Ala Thr Ser Ala Ser Ile Gly Ala Ala Gly Val Pro Gln Ala Gly  
405 410 415

Leu Val Thr Met Val Ile Val Leu Ser Ala Val Gly Leu Pro Ala Glu  
420 425 430

Asp Val Thr Leu Ile Ile Ala Val Asp Trp Leu Leu Asp Arg Phe Arg  
435 440 445

Thr Met Val Asn Val Leu Gly Asp Ala Phe Gly Thr Gly Ile Val Glu  
450 455 460

Lys Leu Ser Lys Lys Glu Leu Glu Gln Met Asp Val Ser Ser Glu Val  
465 470 475 480

Asn Ile Val Asn Pro Phe Ala Leu Glu Ser Thr Ile Leu Asp Asn Glu  
485 490 495

Asp Ser Asp Thr Lys Lys Ser Tyr Val Asn Gly Gly Phe Ala Val Asp  
500 505 510

Lys Ser Asp Thr Ile Ser Phe Thr Gln Thr Ser Gln Phe  
515 520 525

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

56

56  
-57-

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCGGGTACC GCCATGGAGA AGAGCAAC

28

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGCGTCTAGA TCACAGAACC GACTCCTTG

29

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGCGGGTACC AATATGACTA AAAGCAATG

29

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGCGTCTAGA CTACATCTTG GTTCACTG

29

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs

57



57  
-58-

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGCGGGTACC ACCATGGCAT CTACGGAAG

29

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGCGTCTAGA TTATTTCTCA CGTTTCCAAG

30

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGCGGGTACC GCCATGGGGA AACCGGCG

28

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGCGGGATCC CTAGAACTGT GAGGTCTG

28

58